A Bayesian Approach to Nonlinear Diffraction Tomography

Hervé Carfantan and Ali Mohammad-Djafari
Laboratoire des Signaux et Systèmes (CNRS–ESE–UPS)
Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France.
Phone: (33) 1 69 85 17 41, Fax: (33) 1 69 41 30 60 E-mail: carfantan@lss.supelec.fr

Edics Categories: IP 2.3 Tomography, IP 2.1 Acoustic Imaging

I. Problem statement

The reconstruction of some physical properties of an object from the measurement of diffracted waves is a common inverse scattering problem. We set the equations of this problem for the case of electromagnetic waves, in a configuration where the measurement can be performed around the object. However, one can find the same type of equations for acoustic waves and for other configurations, for example in Eddy current nondestructive evaluation [1] and in seismic geophysical exploration [2].

We consider an inhomogeneous object, embedded in a known homogeneous medium, illuminated with a pure Transverse Magnetic (TM) plane wave, and we assume that the polarization effects can be neglected. Moreover, to deal with a two-dimensional problem, we assume the object to be homogeneous along one dimension. The waves satisfy the following scalar integral equation [3], deduced from Maxwell's equations:

$$\phi(\boldsymbol{r}) = \phi_0(\boldsymbol{r}) + \int_D k_0^2 \mathcal{G}(\boldsymbol{r}, \boldsymbol{r}') x(\boldsymbol{r}') \phi(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}',$$

where ϕ and ϕ_0 are respectively the total and the incident field and $\boldsymbol{r}$ denotes a position in $\mathbb{R}^2$. The object is characterized by the complex relative contrast function $x(\boldsymbol{r})$, called the "object" hereafter, defined by $x(\boldsymbol{r}) = k^2(\boldsymbol{r})/k_0^2 - 1$, where $k^2(\boldsymbol{r})$ is the spatially dependant wave number for the inhomogeneous medium and k_0 is the wave number of the background homogeneous medium. The wave number $k(\boldsymbol{r})$ is related to the dielectric permittivity $\epsilon(\boldsymbol{r})$ and the conductivity $\sigma(\boldsymbol{r})$ of the object by: $k^2(\boldsymbol{r}) = \omega^2 \mu_0 \left(\epsilon(\boldsymbol{r}) + j\sigma(\boldsymbol{r})/\omega\right)$, where μ_0 is the magnetic permeability of vacuum and the assumed time dependance is $e^{j\omega t}$. The Green function $\mathcal{G}$ for the homogeneous background medium is $\mathcal{G}(\boldsymbol{r}, \boldsymbol{r}') = \frac{j}{4} H_0^{(1)}\left(k_0 |\boldsymbol{r} - \boldsymbol{r}'|\right)$ where $H_0^{(1)}$ is the first Hankel function of zero order (other functions have to be used for non homogeneous background medium, *e.g.,* for stratified media...).

We consider an imaging system with the geometrical configuration depicted in Fig. (**??**). The measurement sensors are located at points $\boldsymbol{r}_i \in D_{\mathrm{M}}, i = 1 \ldots n_{\mathrm{M}}$, where n_{M} is the number of sensors, distributed on a circle around the object. The investigated inhomogeneous object is illuminated by a plane wave emitted from one sensor while the other sensors measure the scattered field y. The source position is then relocated to another sensor and the operation is repeated. The scattered field y at a measurement point $\boldsymbol{r}_i \in D_{\mathrm{M}}$, corresponding to an incident

field ϕ_0, is given by the coupled equations:

$$y(\boldsymbol{r}_i) = \iint_{D_\mathrm{O}} k_0^2 \mathcal{G}(\boldsymbol{r}_i, \boldsymbol{r}') x(\boldsymbol{r}') \phi(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}', \ \boldsymbol{r}_i \in D_\mathrm{M} \quad (1)$$

$$\phi(\boldsymbol{r}) = \phi_0(\boldsymbol{r}) + \iint_{D_\mathrm{O}} k_0^2 \mathcal{G}(\boldsymbol{r}, \boldsymbol{r}') x(\boldsymbol{r}') \phi(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}', \ \boldsymbol{r} \in D_\mathrm{O} \quad (2)$$

The first equation reflects the propagation of the wave from the source to the measurement area D_M, while the second one corresponds to the propagation in the object area D_O.

The direct problem consists of calculating the measurements $y(\boldsymbol{r}_i), \boldsymbol{r}_i \in D_\mathrm{M}$, corresponding to a given inhomogeneous object $x(\boldsymbol{r}), \boldsymbol{r} \in D_\mathrm{O}$, for the chosen incident field and geometrical configuration. The relation between the object x and the measurements y is nonlinear and implicit but this forward problem is well posed. It cannot generally be solved analytically (except for very particular cases), but several numerical methods can be used (*e.g.* [4]).

We choose to work within an algebraic framework, discretizing the equations (1-2) with a moment method [5]. Consider $\boldsymbol{\phi} \in \mathbb{C}^{n_\mathrm{O}}$ and $\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{O}}$ two column vectors corresponding to the projection of ϕ and x onto a family of functions $(\psi_l(\boldsymbol{r}))_{l \in \{1 \ldots n_\mathrm{O}\}}, \boldsymbol{r} \in D_\mathrm{O}$, known as "basis" functions:

$$x(\boldsymbol{r}) = \sum_l x_l \psi_l(\boldsymbol{r}), \quad \phi(\boldsymbol{r}) = \sum_l \phi_l \psi_l(\boldsymbol{r}).$$

The discretized version of (1-2) are given by their projection respectively onto families of "test" functions $(\rho_i^\mathrm{M}(\boldsymbol{r}))_{i \in \{1 \ldots n_\mathrm{O}\}}, \boldsymbol{r} \in D_\mathrm{M}$ and $(\rho_k^\mathrm{O}(\boldsymbol{r}))_{k \in \{1 \ldots n_\mathrm{M}\}}, \boldsymbol{r} \in D_\mathrm{O}$:

$$y_i = \sum_l x_l \phi_l \iint_{D_\mathrm{M}} \iint_{D_\mathrm{O}} k_0^2 \psi_l(\boldsymbol{r}') \mathcal{G}(\boldsymbol{r}, \boldsymbol{r}') \rho_i^\mathrm{M}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r}' \mathrm{d}\boldsymbol{r},$$

$$\phi_k = (\phi_0)_k + \sum_l x_l \phi_l \iint_{D_\mathrm{O}} \iint_{D_\mathrm{O}} k_0^2 \psi_l(\boldsymbol{r}') \mathcal{G}(\boldsymbol{r}_k, \boldsymbol{r}') \rho_k^\mathrm{O}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r}' \mathrm{d}\boldsymbol{r}.$$

Therefore, the equations can be written formally:

$$\boldsymbol{y} = \boldsymbol{G}_\mathrm{M} \boldsymbol{X} \boldsymbol{\phi}, \quad (3)$$

$$\boldsymbol{\phi} = \boldsymbol{\phi}_0 + \boldsymbol{G}_\mathrm{O} \boldsymbol{X} \boldsymbol{\phi}, \quad (4)$$

where $\boldsymbol{X}$ is a diagonal matrix $(n_\mathrm{O} \times n_\mathrm{O})$ with the components of the vector $\boldsymbol{x}$ as diagonal elements and $\boldsymbol{y} \in \mathbb{C}^{n_\mathrm{M}}, \boldsymbol{\phi}_0 \in \mathbb{C}^{n_\mathrm{O}}$ are vectors representing respectively the measured data and the incident field on the object.

If the functions $\psi_l, \rho_l^\mathrm{O}$ and ρ_l^M are taken to be constant on small square patches, there exist analytical expressions which approximates the elements of $\boldsymbol{G}_\mathrm{O}$ and $\boldsymbol{G}_\mathrm{M}$ [2].

From the algebraic equations (3-4) we obtained a symbolic explicit relation between the unknowns and the data. This can be done by rewriting (4):

$$\boldsymbol{\phi} = (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})^{-1}\,\boldsymbol{\phi}_0,$$

where the matrix $(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})$ is assumed to be invertible. Reporting the total field $\boldsymbol{\phi}$ on the object in (3) gives us an explicit relation for the data $\boldsymbol{y} = \mathcal{A}(\boldsymbol{x})$ with:

$$\mathcal{A}(\boldsymbol{x}) = \boldsymbol{G}_{\mathrm{M}}\boldsymbol{X}\,(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})^{-1}\,\boldsymbol{\phi}_0. \tag{5}$$

Note that the previous relations have been established for only one source location. However, they remain valid in the case of n_{S} sources, up to little changes in the notations (see App. A). In the same way, changes in the notations allows to consider, when necessary, real values variables (see App. B).

The inverse problem, which we are concerned with consist in determining the object $\boldsymbol{x}$ from a given finite set of noisy data $\boldsymbol{y}$. This problem is intrinsically ill-posed. Moreover, note that we can have $n_{\mathrm{O}} \gg n_{\mathrm{M}} \times n_{\mathrm{S}}$ (number of unknown $\gg$ number of data) so that the system of algebraic equations is highly under-determined. Anyway, even if $n_{\mathrm{O}} \ll n_{\mathrm{M}} \times n_{\mathrm{S}}$, the system is ill-conditioned. Thus it has to be regularized to give a satisfactory solution.

II. A Bayesian approach

The Bayesian estimation provides a common framework to deal with inverse problems [6]. It gives a coherent setting to regularize an ill-posed problem by introducing *a priori* information on the solution to fill in the lack of information provided by the data. Such an approach has been used mainly for linear problems. It will help us to define a regularized solution to the nonlinear inverse problem of diffraction tomography, and to compare the different existing methods as algorithms to compute this solution.

To estimate a set of parameters x from experimental data y, it uses Bayes' rule to combine – in the *a posteriori* law – information provided by the data and *a priori* information:

$$p(x \mid y) = \frac{p(y \mid x)p(x)}{p(y)}.$$

$p(y \mid x)$ accounts for both the relation between the unknown x and the data y and the unavoidable uncertainties on the data (errors on the model, discretization errors as well as measurement noise). The *a priori* state of knowledge, that is before any measurement is carried on, is modeled through the probability law $p(x)$. $p(y)$ is a normalization term which does not depend on the parameter x.

The posterior law gather all the information available on x and one can give an estimated value to x considering a punctual estimator such as the mean value (posterior mean, PM), or the maximum value (maximum *a posteriori*, MAP) of this probability law. The MAP estimate is often chosen, as it transforms the estimation problem in an optimization one, and will be considered in the following.

This general framework can be applied on many ways to the nonlinear problem of diffraction tomography, according to the choice of each model. We model the errors on the measurement with an additive, zero mean, white Gaussian circular noise, which seems to be reasonable in lack of additionnal information. The choice of the prior law is a crucial point in the Bayesian framework and is not the aim of this work. For convenience, we take the prior law of the form:

$$p(\boldsymbol{x}) \propto \exp\left\{-\mu\, \mathcal{U}(\boldsymbol{x})\right\},$$

with an energy function $\mathcal{U}$, and the choice of the energy function will be discussed in § II-C.

From these asumptions, we propose two formulations depending on the choice of the parameters to estimate. In the first formulation, only the contrast function $\boldsymbol{x}$ is considered to be estimated from the data, while in the second both $\boldsymbol{x}$ and the field in the object $\boldsymbol{\phi}$ are considered to be estimated.

A. First formulation: Estimation of $\boldsymbol{x}$.

This formulation is straightforward using the explict relation (5) relating the object $\boldsymbol{x}$ and the data $\boldsymbol{y}$. The solution is defined as the MAP estimate of $\boldsymbol{x}$:

$$\boldsymbol{x}_{\text{MAP}} = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x} \,|\, \boldsymbol{y}),$$

which corresponds to the minimizer of the criterion:

$$\mathcal{J}_{\text{MAP}}(\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x})\|^2 + \lambda\, \mathcal{U}(\boldsymbol{x}). \tag{6}$$

B. Second formulation: Joint estimation of $\boldsymbol{x}$ and $\boldsymbol{\phi}$.

The solution is defined as the joint MAP estimate of $\boldsymbol{x}$ and $\boldsymbol{\phi}$:

$$(\boldsymbol{x}, \boldsymbol{\phi})_{\text{MAP}} = \arg\max_{(\boldsymbol{x},\boldsymbol{\phi})} p(\boldsymbol{x}, \boldsymbol{\phi} \,|\, \boldsymbol{y}).$$

Using Bayes' rule, this joint posterior law can be written:

$$p(\boldsymbol{x}, \boldsymbol{\phi} \,|\, \boldsymbol{y}) = \frac{p(\boldsymbol{y} \,|\, \boldsymbol{x}, \boldsymbol{\phi}) p(\boldsymbol{\phi} \,|\, \boldsymbol{x}) p(\boldsymbol{x})}{p(\boldsymbol{y})}.$$

Using (3), with the considered error model, the first term of the numerator can be written:

$$p(\boldsymbol{y} \,|\, \boldsymbol{x}, \boldsymbol{\phi}) \propto \exp\left\{ -\|\boldsymbol{y} - \boldsymbol{G}_{\text{M}} \boldsymbol{X} \boldsymbol{\phi}\|^2 \right\}.$$

The second term corresponds to the probability law of $\boldsymbol{\phi}$ for a known $\boldsymbol{x}$. As $\boldsymbol{\phi}$ is the total field in the object, it is uniquely determined, for a given $\boldsymbol{x}$, by (4). So, if δ denotes the Dirac distribution:

$$p(\boldsymbol{\phi} \,|\, \boldsymbol{x}) = \delta(\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\text{O}} \boldsymbol{X} \boldsymbol{\phi}). \tag{7}$$

Finally, using these expression, the joint posterior probability law can be written:

$$p(\boldsymbol{x}, \boldsymbol{\phi} \,|\, \boldsymbol{y}) \propto \exp\left\{ -\|\boldsymbol{y} - \boldsymbol{G}_{\text{M}} \boldsymbol{X} \boldsymbol{\phi}\|^2 - \lambda\, \mathcal{U}(\boldsymbol{x}) \right\} \delta(\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\text{O}} \boldsymbol{X} \boldsymbol{\phi}).$$

So the joint MAP estimate of $\boldsymbol{x}$ and $\boldsymbol{\phi}$ is the minimizer of the criterion:

$$\mathcal{J}_{\text{JMAP}}(\boldsymbol{x}, \boldsymbol{\phi}) = \|\boldsymbol{y} - \boldsymbol{G}_{\text{M}} \boldsymbol{X} \boldsymbol{\phi}\|^2 + \lambda\, \mathcal{U}(\boldsymbol{x}), \tag{8}$$

subject to the constraint:

$$\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\text{O}} \boldsymbol{X} \boldsymbol{\phi} = \boldsymbol{0}. \tag{9}$$

C. Modeling of prior information

The choice of the energy function $\mathcal{U}$ is essential since it represents prior knowledge on the object to reconstruct. Different considerations, such as energetic or entropic ones [7], [8], can play a part in the modeling of prior information in the regularization theory. However, in this work, we consider the more pragmatic Gibbs-Markov random fields representation, which is frequently used in image processing. It is a flexible way to account for the locally correlated nature of an image by setting probabilistic relationship between neighbor pixels. It leads to an energy function $\mathcal{U}$ of the form:

$$\mathcal{U}(\boldsymbol{x}) = \sum_{c \in \mathcal{C}} V_c(\boldsymbol{x}),$$

where $\mathcal{C}$ is the set of cliques associated to the chosen neighborhood system, and V_c is a potential function on the clique c. The first order neighbors of a site and the associated cliques are illustrated on Fig. **??**. The neighborhood and the potential functions should reflect the prior information on the object.

Many markovian models, or equivalently many potential functions have been used in image processing. Moreover some proposed regularization methods can be understood as Markovian models. For example the classical zeroth order Tikhonov regularization corresponds to a Gaussian prior model ($\mathcal{U}(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$) [6]. One major challenge, is to model an image with smooth areas and discontinuity between the areas. It leads to convex energy functions (*e.g.* [9]) or, by modeling implicit or explicit line processes, to nonconvex energy function (*e.g.* [10], [11], [12]). The choice of such a function is not the goal of this paper and is in itself a wide research area.

Following the work of Bouman and Sauer [9], we chose the Generalized Gauss-Markov model, with the potential function:

$$V_c(\boldsymbol{x}) = |x_j - x_i|^p, \;\; x_j, x_i \in c, \;\; 1 < p \leq 2.$$

For p near or equal 2, it severely penalizes large differences between the values of the neighbors and is useful to reconstruct smoothly varying images. On the other hand, for p near to 1, it penalizes less such differences and could be used to reconstruct objects with sharp transitions. The cliques are taken of first order on horizontal and vertical adjacent pixels. However, all the following development can be made for any other convex energy function $\mathcal{U}$.

The Generalized Gauss-Markov model has been introduced for real value variables and has to be adapted in our case to the complex nature of the variables. In our problem, the real and the

imaginary parts of the object have distinct physical interpretations (respectively the permittivity and the conductivity). For simplicity, we will consider these two parts to be independent and identically distributed. The corresponding energy function can be written:

$$\mathcal{U}_{\mathbb{C}}(\boldsymbol{x}) = \mathcal{U}_{\mathbb{R}}(\mathcal{R}\{\boldsymbol{x}\}) + \mathcal{U}_{\mathbb{R}}(\mathcal{I}\{\boldsymbol{x}\}),$$

where $\mathcal{U}_{\mathbb{R}}$ is the energy function of the Generalized Gauss-Markov model for real value variables.

It should be more realistic to introduce correlation between the real and imaginary parts or more precisely on the discontinuity in these variables, as a discontinuity corresponds to a change in the medium, thus to a simultaneous change of the permittivity and the conductivity of the contrast function. `However, such a choice, is not the subject of this paper....` and the following is valid for any convex energy function $\mathcal{U}$.

D. A calculation challenge

We defined the solution, in the first formulation, as the maximum *a posteriori* estimate of the contrast function $\boldsymbol{x}$, or equivalently as the minimizer of the criterion $\mathcal{J}_{\text{MAP}}$. In the second formulation, the solution is defined as the joint MAP estimate of the contrast function $\boldsymbol{x}$ and the field in the object $\boldsymbol{\phi}$, or equivalently as the joint minimizer of $\mathcal{J}_{\text{JMAP}}$, with respect to $\boldsymbol{x}$ and $\boldsymbol{\phi}$, satisfying the constraint (9). These two distinct formulations are equivalent in the sense they define the same solution (in $\boldsymbol{x}$). However, one may think to use different techniques to solve them.

Due to the non-linearity of $\mathcal{A}$, the data term of $\mathcal{J}_{\text{MAP}}$ is not quadratic with respect to $\boldsymbol{x}$ and it is easy to show that it is not convex. Similarly, the data term of $\mathcal{J}_{\text{JMAP}}$ is globally nonconvex, however, thanks to the bilinearity of Eq. (3), it is convex with respect to each variable $\boldsymbol{x}$ and $\boldsymbol{\phi}$. Due to non-convexity, these criteria may have local minima, even with convex energy functions $\mathcal{U}$. We illustrated this by plotting on Fig. **??** the criterion $\mathcal{J}_{\text{MAP}}$ in two dimensions. This is not, of course, a proof of existence of local minima but it suggest that minimizing such a criterion may be a cumbersome task. Moreover, on some simulation examples, minimizing the criterion $\mathcal{J}_{\text{MAP}}$ with a gradient descent technique, such local minima has been reached (local minimization of $\mathcal{J}_{\text{JMAP}}$ under constraint (9) also leads to such a minima). From simulation experiments, the emergence of such local minima is closely linked to the *difficulty of the configuration*, for example when the contrast is high, when the number of data is very limited or when the data are very noisy. `to illustrate this, ... trace de l'erreur de reconstruction`

`en fonction de la valeur du contraste, Figure these p43` Thus, in this configurations, the minimization of $\mathcal{J}_{\text{MAP}}$ and $\mathcal{J}_{\text{JMAP}}$ should be performed with a global optimization algorithm. However, the problem seems to be less difficult in more favorable configurations.

A large number of methods have been proposed to solve the nonlinear diffraction tomography problem. We propose a classification of most of these methods in terms of algorithm to compute

III. Successive linearizations

A natural idea, when dealing with non linear problems, is to try to account for a linear approximation of its equations. Various linear approximations have been proposed for this inverse scattering problem, among which the well known Born and Rytov approximations. Many methods have been proposed to solve the resulting linear inverse problem, and the limitations of these approximations has been studied has well.

The next idea, is to successively linearize the equations, to transform a nonlinear problem in a sequence of linear ones. For nonlinear inverse problems [13], it leads to consider at each iteration n, a linear approximation of the direct model "*near*" the current solution $\boldsymbol{x}_n$, : $\boldsymbol{y} \approx \boldsymbol{A}_n \boldsymbol{x}$, and to solve the corresponding linear inverse problem. Once again, various methods have been proposed for the inverse scattering problem, from various considerations on the field equations.

Before studying some of these successive linearizations methods, such as the *Born Iterative Method*, the *Distorted Born Iterative Method*, and the *Newton Kantorovitch method*, we will first recall the classical linear approximations. Then we will study the solutions given by these methods in the considered Bayesian framework. Resulting of this study, we propose a coherent successive linearizations algorithm to compute the MAP estimate of the contrast function $\boldsymbol{x}$.

A. Born and first order approximations

The Born approximation considers the diffraction to be weak in the object, substituting ϕ by ϕ_0 in Eq. (1), resulting to a linear model:

$$y(\boldsymbol{r}_i) = \iint_{D_\mathrm{O}} k_0^2 \mathcal{G}(\boldsymbol{r}_i, \boldsymbol{r}') x(\boldsymbol{r}') \phi_0(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}', \ \boldsymbol{r}_i \in D_\mathrm{M}.$$

In the considered algebraic framework, this relation corresponds to the first order Taylor series expansion[1] of $\mathcal{A}(\boldsymbol{x})$ near $\boldsymbol{x} = \boldsymbol{0}$ (see Append. C):

$$\boldsymbol{y} = \boldsymbol{A}_0 \boldsymbol{x} + \mathcal{O}(\boldsymbol{x}^2),$$

with $\boldsymbol{A}_0 = \boldsymbol{G}_\mathrm{M} \boldsymbol{\Phi}_0$, where $\boldsymbol{\Phi}_0$ is the diagonal matrix corresponding to $\boldsymbol{\phi}_0$.

As the Born approximation is a first order approximation on the modulus of ϕ, the Rytov approximation corresponds to a first order approximation on the complex phase of ϕ.

The *distorted wave Born* and *distorted wave Rytov* approximations are Born or Rytov like approximations which consider the object x as a small perturbation δx to a known (and possibly

[1] One should use the real valued notation of App. B to properly use the notion of Taylor series expansion.

strongly scattering) background object x_1. For example, the distorted wave Born approximation corresponds to the Born approximation in which the Green function $\mathcal{G}_1$ accounts for the known inhomogeneous background medium, and the incident field ϕ_0 is substituted by the total field ϕ_1 corresponding to x_1:

$$y(\boldsymbol{r}_i) = \iint_{D_{\mathrm{O}}} k_0^2 \mathcal{G}_1(\boldsymbol{r}_i, \boldsymbol{r}') \delta x(\boldsymbol{r}') \phi_1(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}', \;\; \boldsymbol{r}_i \in D_{\mathrm{M}}.$$

The inverse scattering problem has been considered under the Born or other first order approximations for a long time and many solutions have been proposed. These methods will not be described hereafter has the inverse problem is linear under such an assumption, and linear inverse problems, while still of interest, have been extensively studied in the literature. Note that, accounting for the Born approximation, the considered Bayesian framework leads to minimize the criterion :

$$\mathcal{J}_0(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}_0 \boldsymbol{x}\|^2 + \lambda \mathcal{U}(\boldsymbol{x}),$$

which corresponds to a convex (if the energy function $\mathcal{U}$ is convex) approximation of the criterion $\mathcal{J}_{\mathrm{MAP}}$ near $\boldsymbol{x} = \boldsymbol{0}$. This criterion is unimodal and its minimization can be performed easily, for example with a first order descent technique.

However, it has been shown [14] that these first order approximation fails for large size scatterers or large contrast (with less restrictive conditions for the Rytov than the Born approximation), which is unacceptable for many practical applications. Thus, one has to go beyond these approximations.

B. The Born Iterative Method

B.1 Presentation of the BIM

In [15], inspired by the work of Johnson and Tracy in 1-D [16], Wang and Chew proposed to solve successively and independently each of the coupled equations (3-4). The so called Born Iterative Method (BIM), consists in solving iteratively each of these equations:

1. Equation (3), using a Born type approximation, allows to compute x_{n+1} for a know field on the object ϕ_n (ϕ can be initialized with the incidental field ϕ_0).
2. Equation (4) allows to update the total field ϕ_{n+1} corresponding to the new object x_{n+1}.

Note that the first step corresponds to solve a linear inverse problem, which can be written in the considered algebraic framework:

$$\boldsymbol{y} = \boldsymbol{A}_n^{\mathrm{BIM}} \boldsymbol{x}, \text{ with } \boldsymbol{A}_n^{\mathrm{BIM}} = \boldsymbol{G}_{\mathrm{M}} \boldsymbol{\Phi}_n,$$

where Φ_n is the diagonal matrix corresponding to the field on the object for the given $\boldsymbol{x}_n$: $\boldsymbol{\phi}_n = (I - \boldsymbol{G}_{\mathrm{O}} X_n)^{-1} \boldsymbol{\phi}_0$.

A Tikhonov regularization is used in [15] to regularized each of these problems.

The convergence of the BIM is not guaranteed. However, it has given good practical results, even in cases when the Born approximation is not valid [15], and it has inspired numerous works (*e.g* [17], [18], [19], [20]).

B.2 Analysis of the BIM in the Bayesian framework

In the Bayesian framework, the MAP estimate of $\boldsymbol{x}$ leads to compute the minimizer of criterion $\mathcal{J}_{\mathrm{MAP}}$. It is interesting to analyze the BIM in terms of ability to compute this solution.

As the tikhonov regularization define the solution of a linear inverse problem $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ as the minimizer of $\boldsymbol{J}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda \|\boldsymbol{x}\|^2$, the algorithm corresponding to the BIM can be written on the following way:

TABLE I

THE BORN ITERATIVE METHOD

Set $n = 0$. Initialize $\boldsymbol{x}_0$.

1. Compute the field $\boldsymbol{\phi}_n$ corresponding to the object $\boldsymbol{x}_n$ (*e.g.*, with $\boldsymbol{\phi}_n = (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}} \boldsymbol{X}_n)^{-1} \boldsymbol{\phi}_0$), and the matrix $\boldsymbol{A}_n^{\mathrm{BIM}} = \boldsymbol{G}_{\mathrm{M}} \Phi_n$.
2. Compute $\boldsymbol{x}_{n+1} = \arg\min_{\boldsymbol{x}} \mathcal{J}_n^{\mathrm{BIM}}(\boldsymbol{x})$ with $\mathcal{J}_n^{\mathrm{BIM}}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}_n^{\mathrm{BIM}} \boldsymbol{x}\|^2 + \lambda \mathcal{U}(\boldsymbol{x})$.

Iterate for n on steps 1. and 2. until convergence.

The original BIM corresponding to the particular case $\mathcal{U}(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ and $\boldsymbol{x}_0 = \boldsymbol{0}$.

So, the analysis of BIM in terms of minimization of $\mathcal{J}_{\mathrm{MAP}}$ is straightforward. At each iteration, the criterion $\mathcal{J}_{\mathrm{MAP}}$ is approximated locally by the convex (if the energy function $\mathcal{U}(\boldsymbol{x})$ is convex) criterion $\mathcal{J}_n^{\mathrm{BIM}}$, with $\mathcal{J}_n^{\mathrm{BIM}}(\boldsymbol{x}_n) = \mathcal{J}_{\mathrm{MAP}}(\boldsymbol{x}_n)$ (indeed $\mathcal{A}(\boldsymbol{x}_n) = \boldsymbol{A}_n^{\mathrm{BIM}} \boldsymbol{x}$). Such a scheme is illustrated in 1-D on Fig. III-B.2.

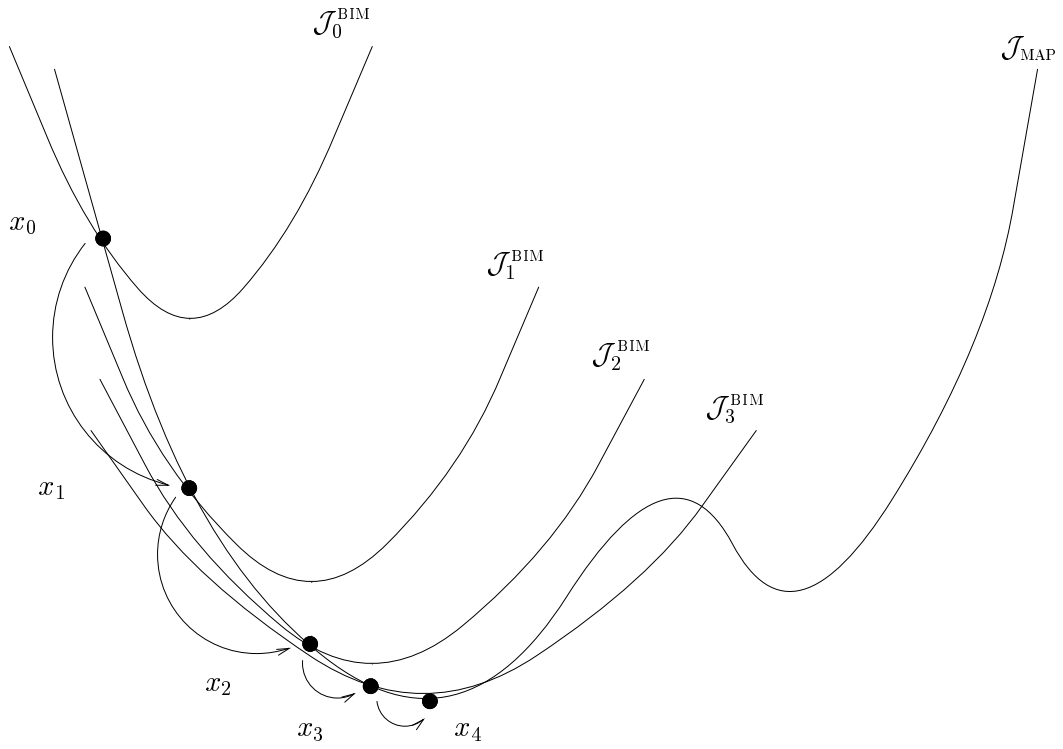

Fig. 1. 1-D illustration of the minimization of $\mathcal{J}_{\mathrm{MAP}}$ using the Born Iterative Method (BIM).

Theoretically, there is no guarantee that this algorithm converges toward a stationary point. Moreover, when it converges towards $\boldsymbol{x}_\infty$:

$$\boldsymbol{x}_\infty = \arg\min_x \left\{ \|\boldsymbol{y} - \boldsymbol{G}_{\mathrm{M}} \boldsymbol{X} (I - \boldsymbol{G}_{\mathrm{O}} \boldsymbol{X}_\infty)^{-1} \boldsymbol{\phi}_0 \|^2 + \lambda \mathcal{U}(\boldsymbol{x}_\infty)) \right\},$$

the gradient of $\mathcal{J}_{\mathrm{MAP}}$ at this point does not necessarily vanishes so that $\boldsymbol{x}_\infty$ does not corresponds to a minimum, even local, of the criterion $\mathcal{J}_{\mathrm{MAP}}$ Thus the BIM is not a satisfactory algorithm to compute $\boldsymbol{x}_{\mathrm{MAP}}$.

C. The Distorted Born Iterative Method

The *distorted Born iterative* method [17] is based on a scheme similar to the BIM, using the Distorted wave Born approximation. At each iteration, an inhomogeneity δx has to be computed, from a known inhomogeneous background medium with contrast x_n, with corresponding Green function $\mathcal{G}_n$ and incident field ϕ_n. Then, the Green function $\mathcal{G}_{n+1}$ and the field on the object ϕ_{n+1} have to be updated for the current inhomogeneous medium $x_{n+1} = x_n + \delta x$. (The equations will not be detailed hereafter.) Once again, the first step corresponds to solve a linear inverse problem and a Tikhonov regularization is used in [17] to compute δx.

The convergence of the DBIM is not guaranteed, furthermore, it has been shown in [17] to diverge in some situations when the BIM was converging. However, the DBIM is still of interest in recent works (*e.g.,* [21]).

D. The Newton-Kantorovitch Method

D.1 Presentation of the NKM

The Born Iterative and Distorted Born iterative methods are specific to the modeling of the direct problem with the coupled equations (3-4). On the other hand, the Newton-Kantorovitch Method [22] (NKM) is a successive linearizations method suitable for every nonlinear inverse problems. To solve nonlinear equations $y = \mathcal{A}(x)$, an iterative scheme is introduced, initialized to x_0, whose iteration consists of computing the variation δx added to x_n so that $y = \mathcal{A}(x_n + \delta x)$. As δx is assumed to be small, $\mathcal{A}(x_n + \delta x)$ is linearized, for each iteration, leading to the linear problem: $y - \mathcal{A}(x_n) = \mathcal{A}_x \delta x$, where $\mathcal{A}_x$ is the Fréchet[2] derivative of $\mathcal{A}$. Of course, the convergence of such a scheme is not guaranteed.

The NKM has been applied to the concerned problem in [23], using a Tikhonov regularization to compute δx. The considered linear approximation of $\mathcal{A}$ can be written in the algebraic framework:

$$\mathcal{A}(x_n + \delta x) = \boldsymbol{A}_n^{\mathrm{NKM}}(x_n + \delta x), \qquad \text{with}^3 \quad \boldsymbol{A}_n^{\mathrm{NKM}} = \boldsymbol{G}_{\mathrm{M}}(\boldsymbol{I} - \boldsymbol{X}_n \boldsymbol{G}_{\mathrm{O}})^{-1} \boldsymbol{\Phi}_n,$$

which corresponds to the first order Taylor series expansion of $\mathcal{A}$ (see Append. C). This approximation has been proven to be equivalent to the DBIM one [24], so that the NKM and the DBIM are equivalent for this problem and only the NKM will be considered hereafter (as it is more general). Compared to the BIM, it can be shown that NKM performs a "*better*" approximation of $\mathcal{A}$ than BIM as $\boldsymbol{A}_n^{\mathrm{BIM}}$ corresponds to a zeroth order approximation of $(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}} \boldsymbol{x})^{-1}$ in the calculation of the first order Taylor series expansion of $\mathcal{A}$ (see Append. C).

D.2 Analysis of the DBIM and NKM in the Bayesian framework

Once again, it is interesting to analyze the NKM (or equivalently the DBIM) as algorithms to compute the MAP estimate of $\boldsymbol{x}$. solution.

The algorithms can be written: where the DBIM of [17] and NKM of [23] corresponds to the

[2]The Fréchet derivative is a generalization of gradient to the infinite dimension space.

[3]Or equivalently $\boldsymbol{A}_n^{\mathrm{NKM}} = \boldsymbol{G}_{\mathrm{M}} \left[\boldsymbol{I} + \boldsymbol{X}_n (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}} \boldsymbol{X}_n)^{-1} \boldsymbol{G}_{\mathrm{O}}\right] \boldsymbol{\Phi}_n$, which can be verified calculating their product.

TABLE II

THE DISTORTED BORN ITERATIVE METHOD AND THE NEWTON-KANTOROVITCH METHOD

Set $n = 0$. Initialize $\boldsymbol{x}_0$.

1. Compute the field $\boldsymbol{\phi}_n$ corresponding to the object $\boldsymbol{x}_n$ (*e.g.*, with $\boldsymbol{\phi}_n = (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{\phi}_0$), and the matrix $\boldsymbol{A}_n^{\mathrm{NKM}} = \boldsymbol{G}_{\mathrm{M}}(\boldsymbol{I} - \boldsymbol{X}_n\boldsymbol{G}_{\mathrm{O}})^{-1}\boldsymbol{\Phi}_n$.
2. Compute $\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \delta\boldsymbol{x}$ with $\delta\boldsymbol{x} = \arg\min_{\delta\boldsymbol{x}} = \mathcal{J}_n^{\mathrm{NKM}}(\delta\boldsymbol{x})$, and $\mathcal{J}_n^{\mathrm{NKM}}(\delta\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_n) - \boldsymbol{A}_n^{\mathrm{NKM}}\delta\boldsymbol{x}\|^2 + \lambda\mathcal{U}(\delta\boldsymbol{x})$.

Iterate for n on steps 1. and 2. until convergence.

particular case $\mathcal{U}(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ and $\boldsymbol{x}_0 = \boldsymbol{0}$.

The analysis of such an algorithm in terms of minimization of $\mathcal{J}_{\mathrm{MAP}}$ is not straightforward because the penalization terms $\mathcal{U}$ operate on the variation $\delta\boldsymbol{x}$ and not on $\boldsymbol{x}$. Everything happen as if each linear inverse problem was regularized independently, without accounting for the fact that they derive from a nonlinear one.

However, this viewpoint allows us to understand the paradox of the DBIM divergence. Indeed, it was noticed in [17] that the DBIM can diverge in configurations where the BIM was still converging. This seems to be contradictory with the fact that the DBIM considers a better approximation of $\mathcal{A}$ than the BIM. However, one may now understand that this unlikely behavior is due to the regularization term in the BIM which only stabilize each linear inverse problem and does not regularize the nonlinear one as a whole. It has been proposed in [23] to account for an adaptative scheme of the regularization parameter λ in the NKM, to overcome such a behavior, and recently, a Modified DBIM as been proposed [21], with the same goal.

E. A Successive linearizations algorithm to minimize $\mathcal{J}_{\mathrm{MAP}}$

From the previous analyze of BIM, DBIM and NKM in terms of minimization of the criterion $\mathcal{J}_{\mathrm{MAP}}$, it appears that none of this methods is satisfactory both in terms of linear approximation and regularization of the nonlinear inverse problem. So it is natural to propose a successive linearizations algorithm (SLA) specifically designed to minimize $\mathcal{J}_{\mathrm{MAP}}$, that is to compute the MAP estimate of the object. At a given step, the first order Taylor series expansion of the operator $\mathcal{A}$, is taken into account to linearily approximate the direct model in the MAP criterion:

$$\mathcal{A}(\boldsymbol{x}) = \mathcal{A}(\boldsymbol{x}_n) + \boldsymbol{A}_n^{\mathrm{SLA}}(\boldsymbol{x} - \boldsymbol{x}_n) + \mathcal{O}\left((\boldsymbol{x} - \boldsymbol{x}_n)^2\right) \text{ with } \boldsymbol{A}_n^{\mathrm{SLA}} = \boldsymbol{A}_n^{\mathrm{NKM}}. \tag{10}$$

This leads to the following algorithm:

TABLE III

THE PROPOSED SUCCESSIVE LINEARIZATION ALGORITHM TO MINIMISE $\mathcal{J}_{\text{MAP}}$

Set $n = 0$. Initialize $\boldsymbol{x}_0$.

1. Compute the field $\boldsymbol{\phi}_n$ and the matrix $\boldsymbol{A}_n^{\text{SLA}}$ corresponding to the linear approximation of $\mathcal{A}$ in the neighborhood of the current solution $\boldsymbol{x}_n$.
2. Compute $\boldsymbol{x}_{n+1} = \arg\min_{\boldsymbol{x}} \mathcal{J}_n^{\text{SLA}}(\boldsymbol{x})$ with $\mathcal{J}_n^{\text{SLA}}(\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_n) - \boldsymbol{A}_n^{\text{SLA}}(\boldsymbol{x} - \boldsymbol{x}_n)\|^2 + \lambda \mathcal{U}(\boldsymbol{x})$.

Iterate for n on steps 1. and 2. until convergence.

At each iteration, the criterion $\mathcal{J}_{\text{MAP}}$ is approximated locally by a convex (if $\mathcal{U}(\boldsymbol{x})$ is convex) criterion $\mathcal{J}_n^{\text{SLA}}$ having the same value for $\boldsymbol{x}_n$ and – due to the first order Taylor series expansion – the same slope at this point (see Append. D). Such a scheme is illustrated in 1-D on Fig. III-E.

As the for BIM, the DBIM and the NKM, there is no guarantee for this algorithm to converge towards a stationary point and the possible convergence depends on the initialization. However, when the SLA converges towards a solution $\boldsymbol{x}_\infty$, it corresponds to a stationnary point of $\mathcal{J}_{\text{MAP}}$, because $\mathcal{J}$ has a null gradient at $\boldsymbol{x}_\infty$ and could hopefully be a minimum, at least local of $\mathcal{J}_{\text{MAP}}$.

One can imagine improving such an algorithm using a second order Taylor series expansion of $\mathcal{J}_{\text{MAP}}$ at each iteration, instead of a first order approximation of $\mathcal{A}$. However, as $\mathcal{J}_{\text{MAP}}$ is not convex, such an approximation may not be convex, and may even be concave in some directions, which will produce the divergence of the resulting algorithm.

F. Simulation results for successive linearizations methods

(Peut etre parler des resultats de simulations : critere non convexe, pas remarque de convergence vers des points selles, soit convergence vers le min dans des configurations favorables pour lesquelles il n'y a pas de min loc, soit divergence ou convergence vers une solution ininteressante pour des configurations difficiles). parler du cout de calcul, et de la possibilite de reduire ce cout (fft)...

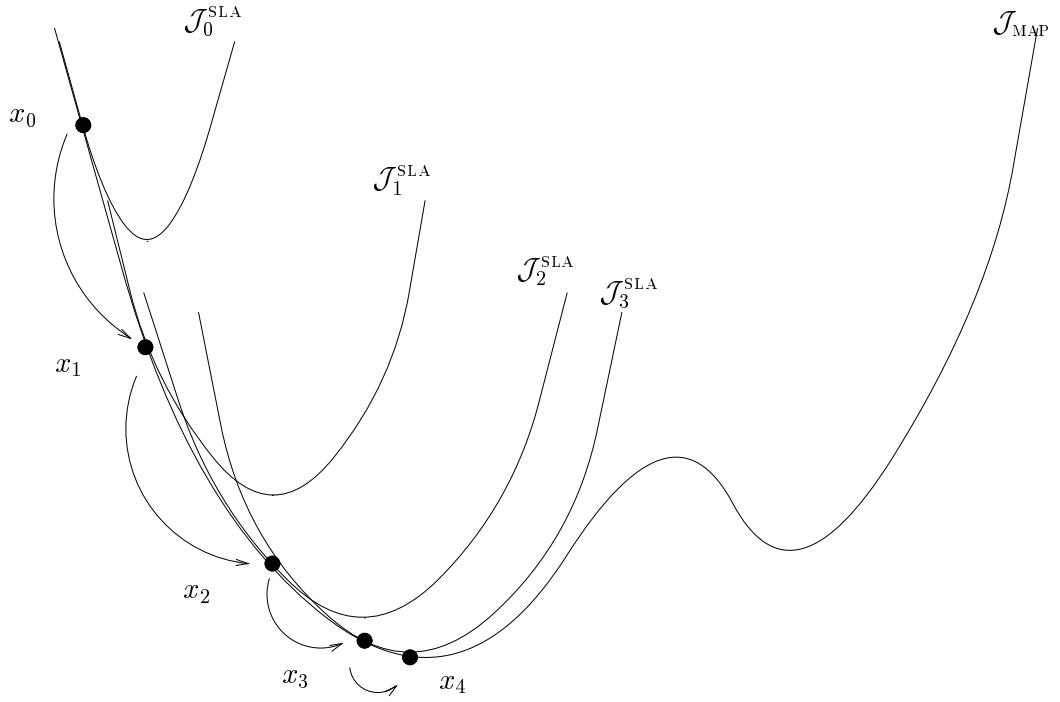

Fig. 2. 1-D illustration of the minimization of $\mathcal{J}_{\mathrm{MAP}}$ using proposed the Successive Linearizations Algorithm (SLA).

IV. Joint Contrast and field computation

Several proposed methods [25], [26], [27] tries to solve this inverse scattering problem computing simultaneously the contrast $\boldsymbol{x}$ and the field in the object $\boldsymbol{\phi}$, from the measured field $\boldsymbol{y}$. The solution is defined as the minimizer of criteria which account for errors on both coupled equations (3-4).

We first briefly present these methods, insisting on their differences. Then we propose a Bayesian interpretation of these methods from the joint formulation of § II-B, (which leads to solve a constrained optimization problem). We establish links between these methods and the computation of the MAP estimate of the contrast function. These links allow us to indicate a coherent way to introduce regularization in these methods, and to analyze their solutions. Finaly, we propose to use a Lagrange multiplier algorithm to solve the constrained optimization problem.

A. Existing methods

These methods define the solution as the contrast $\boldsymbol{x}$ and the field in the object $\boldsymbol{\phi}$, which jointly minimize a criterion of generic form:

$$\boldsymbol{F}(\boldsymbol{x},\boldsymbol{\phi}) = \|\boldsymbol{y} - \boldsymbol{G}_{\mathrm{M}}\boldsymbol{X}\boldsymbol{\phi}\|^2 + \alpha \ \|\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}\boldsymbol{\phi}\|^2 + \lambda \mathcal{U}(\boldsymbol{x},\boldsymbol{\phi}). \tag{11}$$

Such a definition is very easy to understand intuitively: it corresponds to jointly minimizing the errors on (3) and (4) and, as the problem is ill-posed, a regularizing penalization term on the unknowns is added.

The proposed methods differ on several points:

- Different values has been proposed for parameter α. In [25], this parameter is fixed to normalize the errors on both equations for $\boldsymbol{\phi} = \boldsymbol{0}$: $\alpha = \|\boldsymbol{y}\|^2/\|\boldsymbol{\phi}_0\|^2$, while it is fixed to 1 in [26].
- Differences appear on the regularizing term.
  - First, no regularization was introduced [25], [26].
  - Then, it has been proposed to regularize both on $\boldsymbol{x}$ and $\boldsymbol{\phi}$, with an energy function $\mathcal{U}(\boldsymbol{x},\boldsymbol{\phi})$ [27], [28].
  - Finally, a single regularization term on $\boldsymbol{x}$ was accounted for [29], [30].
- The methods also differ in the techniques used to compute the solution. Usual gradient type

local minimization techniques have been used [26], [28] as well as local techniques specially designed for such a criterion [25] and global minimization techniques such as Simulated Annealing [29].

Note that such a method never requires solution of the direct problem – which was one of their main objectives – so it is of relative low computation cost. On the other hand, the number of unknown is multiplied by $n_{\mathrm{S}} + 1$, as the object $\boldsymbol{x}$ and the field in the object for each incident wave have to be determinate.

B. Bayesian viewpoint

Recall from § II-B that joint estimation of $\boldsymbol{x}$ and $\boldsymbol{\phi}$ leads to minimize $\mathcal{J}_{\mathrm{JMAP}}$ of Eq. (8) subject to constraint (9). In this framework, criterion (11) can be understood as the Lagrangian – with a fixed Lagrange parameter α – of this constrained optimization problem (for a scalar constraint: $\|\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}\boldsymbol{\phi}\|^2 = 0$); or has a penalization of $\mathcal{J}_{\mathrm{JMAP}}$ with the error on the constraint. In both cases, the choice of parameter α is important: on one hand, the value of α should be high enough to enforce the constraint; on the other hand, the criterion may become numerically insensitive to the data if α is too high. Note that α has been fixed intuitively in the different methods.

Moreover, this viewpoint gives indications for regularizing such a criterion with an energy function $\mathcal{U}(\boldsymbol{x})$. Using Bayes rule for the considered model of errors on measurements, we can see from Eq. (II-B) that there is no need to introduce prior model on $\boldsymbol{\phi}$.

Another Bayesian interpretation of this criterion has been given in [29]. It has been proposed to account for additive Gaussian models for errors on Eq. (4), so that the joint MAP estimate of $\boldsymbol{x}$ and $\boldsymbol{\phi}$ minimizes a criterion of form (11). That for, the authors "assumed that":

$$p(\boldsymbol{\phi} \,|\, \boldsymbol{x}) \propto \exp\left\{-\|\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}\boldsymbol{\phi}\|^2\right\}.$$

However, it can be shown [31] that to account for such a relation, the error on Eq. (4) should be zero mean circular Gaussian – conditionally to the knowledge of $\boldsymbol{x}$ – with covariance matrix[4]: $C_{\mathrm{O}} = \sigma_2^2 \left[(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})^\dagger(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})\right]^{-1}$.

This seems to be a very strong and unjustified hypothesis as it considers a particular correlation between the errors and the unknowns.

[4]Up to an abuse in notation for complex valued variables.

Whatever their interpretation may be, none of these methods try to compute the MAP estimate has defined in § II-B, but to jointly minimze criterion (11). However, it is easy to show that if $(\boldsymbol{x}, \boldsymbol{\phi})$ is a local minimum of (6) and that constraint (9) is satisfied, then $\boldsymbol{x}$ corresponds to a stationnary point of the unconstrained criterion $\mathcal{J}_{\text{MAP}}$, and is a serious candidate to be the MAP estimate. But nothing special is performed to satisfy te constraint, that is why we propose to use a Lagrange multiplier algorithm to solve this constrained optimization problem.

C. Lagrange multiplier algorithm

This algorithm is an application of the *method of multipliers* [32] – which is based on the augmented Lagrangian – and consists of a synthesis of Lagrangian and penalyzing methods (see [32] for a general presentation of the algorithm and for some indications on the values of the different parameters.)

This leads to the following algorithm [5] [31]:

TABLE IV

THE PROPOSED LAGRANGE MULTIPLIER ALGORITHM TO MINIMISE $\mathcal{J}_{\text{JMAP}}$ UNDER CONSTRAINT.

Initialize $n = 0, \boldsymbol{\mu}_0 = \mathbf{0}, \alpha_0$.

1. Compute $(\boldsymbol{x}_n, \boldsymbol{\phi}_n)$ minimizing the augmented Lagrangian: $\mathcal{L}_{\alpha_n}((\boldsymbol{x}, \boldsymbol{\phi}), \boldsymbol{\mu}_n) = \|\boldsymbol{y} - \boldsymbol{G}_{\text{M}}\boldsymbol{X}\boldsymbol{\phi}\|^2 + \boldsymbol{\mu}_n^{\text{t}}(\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\text{O}}\boldsymbol{X}\boldsymbol{\phi}) + \alpha_n\|\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\text{O}}\boldsymbol{X}\boldsymbol{\phi}\|^2 + \lambda\mathcal{U}(\boldsymbol{x})$;
2. Update parameters: $\alpha_{n+1} = \beta\alpha_n$ (β fixed) and $\boldsymbol{\mu}_{n+1} = \boldsymbol{\mu}_n + \alpha_n(\boldsymbol{\phi} - \boldsymbol{\phi}_0 - \boldsymbol{G}_{\text{O}}\boldsymbol{X}\boldsymbol{\phi})$.

Iterate for n on steps 1. and 2. until convergence.

We used a gradient descent technique to perform the minimization of the augmented Lagrangian, so the reached minimum at each step is only a local minimum, (the user's prefered local minimization algorithm could be used). Such an algorithm, when converges, guarantees to reach a local minimum of (8) which verifies the constraint (9), but unfortunately, the convergence is not guaranteed. However, the first step of this algorithm corresponds to the minimization of (11), and the use of additional iterations has been shown on simulations to decrease the error on the constraint and to improve the solution.

[5] For simplicity of presentation, we present the algorithm in the case of the real valued variables notation of App. B.

D. Simulation results for joint Contrast and field computation methods

Amelioration de l'algo des multiplicateurs de Lagrange: Valeur du critere, erreur sur la contriantes, erreur sur l'objet: figure et tableau. Parler du cout de calcul.

V. Direct minimization of $\mathcal{J}_{\text{MAP}}$

Various authors proposed to minimize the least square criterion, possibly taking into account a regularization term (*e.g.* [33], [34], [35]); which correspond to directly minimize the MAP criterion (6). The computation of this criterion is very expensive as it requires solution of the direct problem. Thus, two cases should be distinguished:

- In *easy configurations*, when $\mathcal{J}_{\text{MAP}}$ – even nonconvex – seems to have a unique minimum, a local optimization technique could be used [34], [35]. However, we have shown that in this case the solution can be computed using a less computationally expensive successive linearizations algorithm (SLA § III-E) or a Lagrange multipliers algorithm (LMA § IV-C); so one should avoid minimizing directly the MAP criterion.
- In *difficult configurations*, when local minima exist, the previously studied methods of § III and § IV fail. In this case, a global optimization technique could be used to minimize $\mathcal{J}_{\text{MAP}}$. Simulated Annealing has been used in [33], but such a technique is practically inextricable because of the high computation cost of the criterion and the large support of the operator $\mathcal{A}$. Instead, we proposed two less expensive deterministic algorithms to try to reach the global minimum:
  - The first algorithm [36], is based on a Graduated Non Convexity (GNC) relaxation scheme.
  - The second one [37] is an Iterated Conditional Mode (ICM) algorithm which performs the minimization by iteratively updating each pixel of the object.

  These algorithms will not be detailed hereafter, but their major principles will be explained.

A. GNC-based algorithm

The GNC principle is quite simple [12]. To compute the global minimum of a criterion $\mathcal{J}$, it consists of building a sequence of criteria $\{\mathcal{J}_n^{\text{GNC}}\}_n$ simply converging toward it (*i.e.*, $\lim_{n\to\infty} \mathcal{J}_n^{\text{GNC}}(\boldsymbol{x}) = \mathcal{J}(\boldsymbol{x}), \forall \boldsymbol{x}$), such that the first of these criteria $\mathcal{J}_0^{\text{GNC}}$ is convex. Then, a local optimization of each criteria $\mathcal{J}_n^{\text{GNC}}$ is performed in the neighborhood of the solution $\boldsymbol{x}_{n-1}$ corresponding to the minimum of the previous criterion $\mathcal{J}_{n-1}^{\text{GNC}}$. To compare with the successive local approximations algorithm of § III (see Fig. (III-B.2) and (III-E)) such an algorithm can be qualified of successive global approximations, as illustrated on Fig. V-A. There is no convergence guarantee of a GNC algorithm towards the global minimum, but at least a guarantee to reach a local minimum which do not depend on the initialization. However, such an algorithm has given

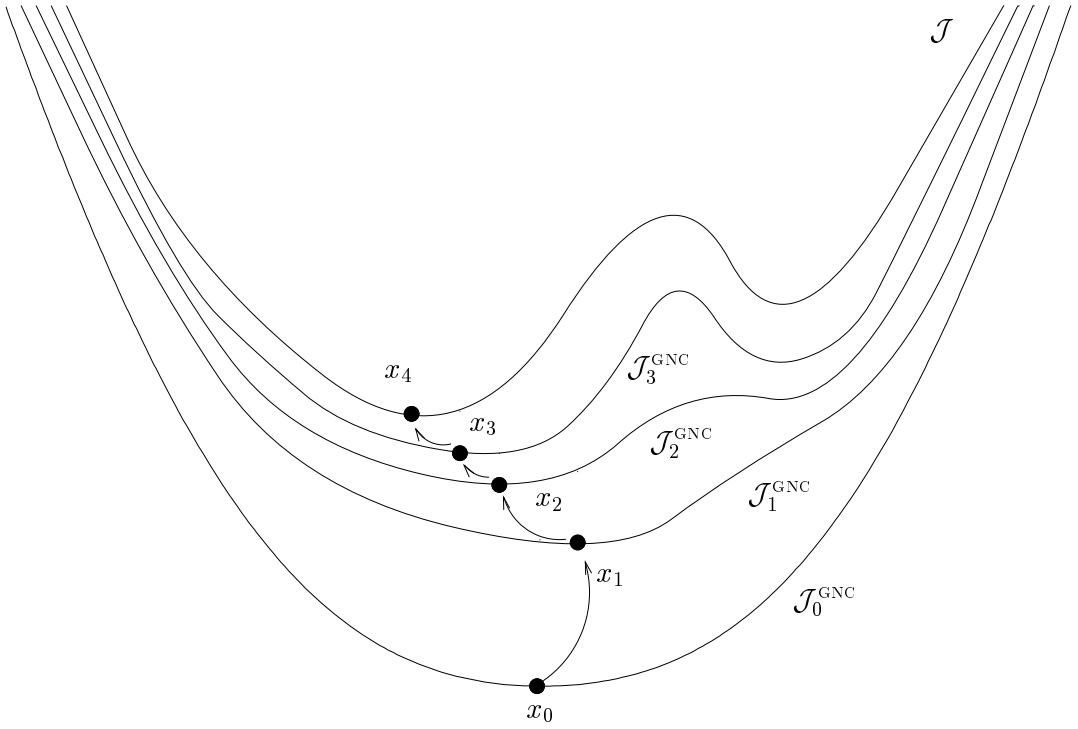

Fig. 3. 1-D illustration of the minimization of a criterion $\mathcal{J}$ using a Graduated Non Convexity (GNC) relaxation scheme.

very interesting results in denoising and segmentation [12] and for various linear inverse problem (*e.g.*, [38]).

Unfortunately, no indication is given to build the relaxation sequence $\{\mathcal{J}_n^{\mathrm{GNC}}\}_n$. In our case – in opposition to the previous applications for which the non-convexity is provided by a nonconvex penalyzing term of the criterion – the regularizing term is convex and the non-convexity of the criterion is due to the non-linearity of the direct problem. So we proposed [36] to operate on this non-linearity to introduce the following sequence of criteria:

$$\mathcal{J}_n^{\mathrm{GNC}}(\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}_n^{\mathrm{GNC}}(\boldsymbol{x})\|^2 + \lambda \mathcal{U}(\boldsymbol{x}),$$

with

$$\mathcal{A}_n^{\mathrm{GNC}}(\boldsymbol{x}) = \boldsymbol{G}_{\mathrm{M}} \boldsymbol{X} (\boldsymbol{I} - r_n \boldsymbol{G}_{\mathrm{O}} \boldsymbol{X})^{-1} \boldsymbol{\phi}_0,$$

where $\{r_n\}$ is a sequence of real relaxation parameters such that $r_0 = 0$, and $\lim_{n\to\infty} r_n = 1$ (in practice, a finite number of iterations $n_{\max}$ is fixed and $r_{n_{\max}} = 1$.) This leads to the following algorithm:

TABLE V

THE PROPOSED GNC BASED ALGORITHM TO MINIMISE $\mathcal{J}_{\text{MAP}}$

Initialisation $n = 0,\ r_0 = 0, \boldsymbol{x}_{-1}$.

Compute $\boldsymbol{x}_n$, local minimum of $\mathcal{J}_n^{\text{GNC}}(\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}_n(\boldsymbol{x})\|^2 + \lambda \mathcal{U}(\boldsymbol{x})$ in the neigborhood of $\boldsymbol{x}_{n-1}$, with $\mathcal{A}_n(\boldsymbol{x}) = \boldsymbol{G}_{\text{M}} \boldsymbol{X} (\boldsymbol{I} - r_n \boldsymbol{G}_{\text{O}} X)^{-1} \boldsymbol{\phi}_0$

Iteration for $n = 1, 2 \ldots n_{\max}$.

Note that for $r_0 = 0$, operator $\mathcal{A}_0^{\text{GNC}}$ correspond to the (linear) Born approximation of the direct problem (see § III-A) and the criterion $\mathcal{J}_0^{\text{GNC}}$, is convex.

`Parler du cout de calcul, illustration du schema de relaxation sur le critere, resultats satisfaisants, meme quand minimum global`

B. ICM algorithm

The second one is an Iterated Conditional Mode (ICM) algorithm which performs the minimization by iteratively updating each pixel of the object. Such an algorithm can be efficiently implemented [37] using special properties of the criterion (6).

None of these algorithms are guaranteed to converge towards the global minimum but both have given satisfactory results in configurations where a gradient type algorithm gets stuck in a local minimum.

Appendix

A. Direct problem model in the multi-sources case.

To take into account all the measurements at the same time, we can consider the matrices $\boldsymbol{y} = \left[\boldsymbol{y}^1, \boldsymbol{y}^2, \ldots, \boldsymbol{y}^{n_\mathrm{S}}\right]$, $\boldsymbol{\phi} = \left[\boldsymbol{\phi}^1, \boldsymbol{\phi}^2, \ldots, \boldsymbol{\phi}^{n_\mathrm{S}}\right]$ and $\boldsymbol{\phi}_0 = \left[\boldsymbol{\phi}_0^1, \boldsymbol{\phi}_0^2, \ldots, \boldsymbol{\phi}_0^{n_\mathrm{S}}\right]$, where $\boldsymbol{y}^i$, $\boldsymbol{\phi}^i$ and $\boldsymbol{\phi}_0^i$ are respectively the measurements, the total and the incident field on the scatterer, corresponding to the i^{th} illumination, and n_S is the number of sources. Therefore, $\boldsymbol{y}$ is an $n_\mathrm{M} \times n_\mathrm{S}$ matrix and $\boldsymbol{\phi}$, $\boldsymbol{\phi}_0$ are $n_\mathrm{O} \times n_\mathrm{S}$. It is interesting to notice that for the chosen configuration, and with the taken notations, the relations verified by the matrices $\boldsymbol{y}$, $\boldsymbol{\phi}$ and $\boldsymbol{\phi}_0$ are strictly the same as those verified in the unique source case, that is Eq. (3–4). Thus, we do not have to differentiate between the cases of a unique or multi-sources for the considered direct model.

However, to solve the inverse problem, the notations used in this paper have to be extended in the multi-sources case:

- We use the usual L_2 norm on the vectors $\boldsymbol{y}$: $\|\boldsymbol{y}\|^2 = \sum_l |\boldsymbol{y}_l|^2$. In the multi-source case, the norm we apply on matrix $\boldsymbol{y}$ has to be understood as: $\|\boldsymbol{y}\|^2 = \sum_{i=1}^{n_\mathrm{S}} \|\boldsymbol{y}^i\|^2$ where $\boldsymbol{y}^i$ is the i^{th} column of $\boldsymbol{y}$.
- In the multisource case $\boldsymbol{\phi}$ and $\boldsymbol{\phi}_0$ are $n_\mathrm{O} \times n_\mathrm{S}$ matrices. Thus we can not define corresponding diagonal matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Phi}_0$. However, we can keep this notation for sake of simplicity, reminding that the product $\boldsymbol{\Phi}\boldsymbol{x}$ (resp. $\boldsymbol{\Phi}_0\boldsymbol{x}$) can be calculated with $X\boldsymbol{\phi}_0$ (resp. $X\boldsymbol{\phi}_0$).

B. Real Valued variable notations

Some quantites manipulated in this paper are only defined for real valued functions of real value variables. A simple extension of the notations allows to carry on with Eq. (3–4) in this real values case. That for, consider vectors:

$$\widetilde{\boldsymbol{x}} = \begin{bmatrix} \Re\{\boldsymbol{x}\} \\ \Im\{\boldsymbol{x}\} \end{bmatrix}, \widetilde{\boldsymbol{y}} = \begin{bmatrix} \Re\{\boldsymbol{y}\} \\ \Im\{\boldsymbol{y}\} \end{bmatrix}, \widetilde{\boldsymbol{\phi}} = \begin{bmatrix} \Re\{\boldsymbol{\phi}\} \\ \Im\{\boldsymbol{\phi}\} \end{bmatrix}, \widetilde{\boldsymbol{\phi}_0} = \begin{bmatrix} \Re\{\boldsymbol{\phi}_0\} \\ \Im\{\boldsymbol{\phi}_0\} \end{bmatrix}$$

and matrices:

$$\widetilde{\boldsymbol{X}} = \begin{bmatrix} \Re\{\boldsymbol{X}\} & -\Im\{\boldsymbol{X}\} \\ \Im\{\boldsymbol{X}\} & \Re\{\boldsymbol{X}\} \end{bmatrix}, \widetilde{\boldsymbol{G}_\mathrm{O}} = \begin{bmatrix} \Re\{\boldsymbol{G}_\mathrm{O}\} & -\Im\{\boldsymbol{G}_\mathrm{O}\} \\ \Im\{\boldsymbol{G}_\mathrm{O}\} & \Re\{\boldsymbol{G}_\mathrm{O}\} \end{bmatrix}, \widetilde{\boldsymbol{G}_\mathrm{M}} = \begin{bmatrix} \Re\{\boldsymbol{G}_\mathrm{M}\} & -\Im\{\boldsymbol{G}_\mathrm{M}\} \\ \Im\{\boldsymbol{G}_\mathrm{M}\} & \Re\{\boldsymbol{G}_\mathrm{M}\} \end{bmatrix}.$$

C. First order Taylor series expansion of the operator $\mathcal{A}$

We have to calculate the first order Taylor series expansion[6] of the operator $\mathcal{A}$ at $\boldsymbol{x} \approx \boldsymbol{x}_n$:

$$\mathcal{A}(\boldsymbol{x}) = \mathcal{A}(\boldsymbol{x}_n) + \nabla\mathcal{A}(\boldsymbol{x}_n)(\boldsymbol{x} - \boldsymbol{x}_n) + \mathcal{O}\left((\boldsymbol{x} - \boldsymbol{x}_n)^2\right).$$

That is to calculate the matrix $\boldsymbol{A}_n^{\mathrm{SLA}} = \nabla\mathcal{A}(\boldsymbol{x}_n)$ such that for small $\delta\boldsymbol{x}$:

$$\mathcal{A}(\boldsymbol{x}_n + \delta\boldsymbol{x}) = \mathcal{A}(\boldsymbol{x}_n) + \boldsymbol{A}_n^{\mathrm{SLA}}\delta\boldsymbol{x} + \mathcal{O}\left((\delta\boldsymbol{x})^2\right),$$

[6] The real valued notation is considered in the following.

from the relation:

$$\mathcal{A}(\boldsymbol{x}_n + \delta\boldsymbol{x}) = \boldsymbol{G}_{\mathrm{M}}(\boldsymbol{X}_n + \delta\boldsymbol{X})\left[\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}(\boldsymbol{X}_n + \delta\boldsymbol{X})\right]^{-1}\boldsymbol{\phi}_0. \tag{12}$$

An elegant manner to get further in this development and to calculate the first order approximation of the matrix $[\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}(\boldsymbol{X}_n + \delta\boldsymbol{X})]^{-1}$ is to use the matrix inversion lemma:

$$(\boldsymbol{P} + \boldsymbol{QRS})^{-1} = \boldsymbol{P}^{-1} - \boldsymbol{P}^{-1}\boldsymbol{Q}(\boldsymbol{R}^{-1} + \boldsymbol{SP}^{-1}\boldsymbol{Q})^{-1}\boldsymbol{SP}^{-1},$$

where the indicated inverses exist. We take $\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)$, $\boldsymbol{Q} = -\boldsymbol{G}_{\mathrm{O}}$, $\boldsymbol{R} = \boldsymbol{I}$ and $\boldsymbol{S} = \delta\boldsymbol{X}$; then

$$\begin{aligned}[(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n) - \boldsymbol{G}_{\mathrm{O}}\delta\boldsymbol{X}]^{-1} &= (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1} \\ &+(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{G}_{\mathrm{O}}\left[\boldsymbol{I} - \delta\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{G}_{\mathrm{O}}\right]^{-1}\delta\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1},\end{aligned}$$

which can simply be approximated to first order with:

$$[(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n) - \boldsymbol{G}_{\mathrm{O}}\delta\boldsymbol{X})]^{-1} = (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\left[\boldsymbol{I} + \boldsymbol{G}_{\mathrm{O}}\delta\boldsymbol{X}(I - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\right] + \mathcal{O}\left((\delta\boldsymbol{x})^2\right). \tag{13}$$

A first order approximation of $\mathcal{A}(\boldsymbol{x}_n + \delta\boldsymbol{x})$ can then be written as:

$$\begin{aligned}\mathcal{A}(\boldsymbol{x}_n + \delta\boldsymbol{x}) &= \boldsymbol{G}_{\mathrm{M}}(\boldsymbol{X}_n + \delta\boldsymbol{X})(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\left[\boldsymbol{I} + \boldsymbol{G}_{\mathrm{O}}\delta\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\right]\boldsymbol{\phi}_0 + \mathcal{O}\left((\delta\boldsymbol{x})^2\right) \\ &= \mathcal{A}(\boldsymbol{x}_n) + \boldsymbol{G}_{\mathrm{M}}\delta\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{\phi}_0 \\ &\quad +\boldsymbol{G}_{\mathrm{M}}\boldsymbol{X}_n(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{G}_{\mathrm{O}}\delta\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{\phi}_0 + \mathcal{O}\left((\delta\boldsymbol{x})^2\right)\end{aligned}$$

Note that if instead of (13), only the matrix $(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}$ is taken into account (zeroth order approximation), the approximation of $\mathcal{A}(\boldsymbol{x}_n + \delta\boldsymbol{x})$ turns to the approximation taken into account in the BIM.

If $\boldsymbol{\phi}_n = (\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{\phi}_0$ denotes the field on the object for the given $\boldsymbol{x}_n$, and Φ_n its associated diagonal matrix , $\boldsymbol{A}_n^{\mathrm{SLA}}$ can be written

$$\boldsymbol{A}_n^{\mathrm{SLA}} = \boldsymbol{G}_{\mathrm{M}}\left[\boldsymbol{I} + \boldsymbol{X}_n(\boldsymbol{I} - \boldsymbol{G}_{\mathrm{O}}\boldsymbol{X}_n)^{-1}\boldsymbol{G}_{\mathrm{O}}\right]\Phi_n.$$

Note that at $\boldsymbol{x}_0 = 0$, $\boldsymbol{A}_0^{\mathrm{SLA}} = \boldsymbol{G}_{\mathrm{M}}\Phi_0$, thus $\mathcal{A}(\boldsymbol{x}) \approx \boldsymbol{G}_{\mathrm{M}}\Phi_0\boldsymbol{x}$, for small $\boldsymbol{x}$ which is equivalent to the Born approximation (strictly speaking: taking $\phi = \phi_0$ instead of Eq. (2), see § (III-A))).

D. Gradient of the MAP and of the successive linearizations criteria

The criterion to minimize at each step in the successive linearizations method is:

$$\mathcal{J}_n^{\mathrm{SLA}}(\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_n) - \boldsymbol{A}_n^{\mathrm{SLA}}(\boldsymbol{x} - \boldsymbol{x}_n)\|^2 + \lambda\mathcal{U}(\boldsymbol{x}).$$

Its gradient can be expressed:

$$\nabla\mathcal{J}_n^{\mathrm{SLA}}(\boldsymbol{x}) = -2(\boldsymbol{A}_n^{\mathrm{SLA}})'(\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_n) - \boldsymbol{A}_n^{\mathrm{SLA}}(\boldsymbol{x} - \boldsymbol{x}_n)) + \lambda\nabla\mathcal{U}(\boldsymbol{x}).$$

Thus, at $\boldsymbol{x} = \boldsymbol{x}_n$:

$$\nabla \mathcal{J}_n^{\mathrm{SLA}}(\boldsymbol{x}_n) = -2(\boldsymbol{A}_n^{\mathrm{SLA}})'(\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_n)) + \lambda \nabla \mathcal{U}(\boldsymbol{x}_n).$$

At $\boldsymbol{x} = \boldsymbol{x}_n$ the gradient of the MAP criterion $JMAP$ is:

$$\nabla \mathcal{J}_{\mathrm{MAP}}(\boldsymbol{x}_n) = -2(\nabla \mathcal{A}(\boldsymbol{x}_n))'(\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_n)) + \lambda \nabla \mathcal{U}(\boldsymbol{x}_n).$$

By construction of $\boldsymbol{A}_n^{\mathrm{SLA}}$ (see Eq. (10): $\boldsymbol{A}_n^{\mathrm{SLA}} = \nabla \mathcal{A}(\boldsymbol{x}_n)$, then we have

$$\nabla \mathcal{J}_{\mathrm{MAP}}(\boldsymbol{x}_n) = \nabla \mathcal{J}_n^{\mathrm{SLA}}(\boldsymbol{x}_n),$$

and the two criteria have the same value and the same slope for $\boldsymbol{x} = \boldsymbol{x}_n$.

References

[1] D. Prémel and A. Mohammad-Djafari, "Eddy current tomography in cylindrical geometry", *IEEE Transactions on Magnetics*, vol. M-31, no. 3, pp. 2000–2003, May 1995.

[2] A. Q. J. Howard and J. L. Kretzschmar, "Synthesis of EM geophysical tomographic data", *Proceedings of the IEEE*, vol. 74, no. 2, pp. 353–360, February 1986.

[3] D. Colton and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, Springer-Verlag, 1992.

[4] R. E. Kleinman and P. M. van den Berg, "Iterative method for solving integral equations", *Radio Science*, vol. 26, no. 1, pp. 175–181, January–February 1991.

[5] R. F. Harrington, "The method of moments in electromagnetics", *Journal of Electromagnetic Waves and Applications*, vol. 1, no. 3, pp. 181–200, March 1987.

[6] G. Demoment, "Image reconstruction and restoration : Overview of common estimation structure and problems", *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. ASSP-37, no. 12, pp. 2024–2036, December 1989.

[7] E. T. Jaynes, "Prior probabilities", *IEEE Transactions on Systems Science and Cybernetics*, vol. SSC-4, no. 3, pp. 227–241, September 1968.

[8] A. Mohammad-Djafari and G. Demoment, "Image restoration and reconstruction using entropy as a regularization functional", in *Maximum Entropy and Bayesian Methods in Science and Engineering*, G. J. Erikson and S. C. Ray, Eds., Dordrecht, The Netherlands, 1988, MaxEnt Workshops, vol. 2, pp. 341–355, Kluwer Academic Publishers.

[9] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation", *IEEE Transactions on Image Processing*, vol. IP-2, no. 3, pp. 296–310, July 1993.

[10] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, November 1984.

[11] D. Geman and G. Reynolds, "Constrained restoration and recovery of discontinuities", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-14, no. 3, pp. 367–383, March 1992.

[12] A. Blake and A. Zisserman, *Visual reconstruction*, The MIT Press, Cambridge, 1987.

[13] A. Tarantola, *Inverse problem theory : Methods for data fitting and model parameter estimation*, Elsevier Science Publishers, Amsterdam, 1987.

[14] M. Slaney, A. C. Kak, and L. E. Larsen, "Limitations of imaging with first-order diffraction tomography", *IEEE Transactions on Microwave Theory and Technology*, vol. MTT-32, no. 8, pp. 860–874, August 1984.

[15] Y. M. Wang and W. C. Chew, "An iterative solution of the two-dimensional electromagnetic inverse scattering problem", *International Journal of Imaging Systems and Technology*, vol. 1, pp. 100–108, 1989.

[16] S. A. Johnson and M. L. Tracy, "Inverse scattering solutions by a sinc basis, multiple source, moment method – part 1: theory", *Ultrasonic Imaging*, vol. 5, pp. 361–375, 1983.

[17] W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method", *IEEE Transactions on Medical Imaging*, vol. MI-9, pp. 218–225, June 1990.

[18] M. Moghaddam, W. C. Chew, and M. L. Oristaglio, "Comparison of the Born iterative method and Tarantola's method for an electromagnetic time-domain inverse problem", *International Journal of Imaging Systems and Technology*, vol. 3, pp. 318–333, 1991.

[19] Q.-H. Liu, "Reconstruction of two-dimensional axisymetric inhomogeneous media", *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-31, no. 3, pp. 587–594, May 1993.

[20] P. Chaturvedi and R. G. Plumb, "Electromagnetic imaging of underground targets using constrained optimization", *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-33, no. 3, pp. 551–561, May 1995.

[21] O. S. Haddadin and E. S. Ebbini, "Adaptative regularisation of a distorted Born iterative algorithm for diffraction tomography", in *Proceedings of the International Conference on Image Processing*, Lausanne, Switzerland, September 1996, pp. 725–728.

[22] A. Roger, "Newton-Kantorovitch algorithm applied to an electromagnetic inverse problem", *IEEE Transactions on Antennas and Propagation*, vol. AP-29, no. 2, pp. 232–238, March 1981.

[23] N. Joachimovicz, C. Pichot, and J.-P. Hugonin, "Inverse scattering: An iterative numerical method for electromagnetic imaging", *IEEE Transactions on Antennas and Propagation*, vol. AP-39, no. 12, pp. 1742–1752, December 1991.

[24] A. Franchois, *Contribution a la tomographie microonde : algorithmes de reconstruction quantitative et vérifications expérimentales*, PhD thesis, Université de Paris-XI, June 1993.

[25] R. E. Kleinman and P. M. van den Berg, "A modified gradient method for two-dimensional problems in tomography", *Journal of Computational and Applied Mathematics*, vol. 42, pp. 17–35, 1992.

[26] H. A. Sabbagh and R. G. Lautzenheiser, "Inverse problems in electromagnetic nondestructive evaluation", *International Journal of Applied Electromagnetics in Materials*, vol. 3, pp. 235–261, 1993.

[27] S. Caorsi, G. L. Gragnani, M. Pastorino, and A. Peraso, "Electromagnetic inverse scattering numerical method for non invasive diagnostic of dielectric materials", in *3rd International Conference on Electromagnetics in Aerospace Applications*, Torino, Italy, September 1993.

[28] S. Barkeshli and R. G. Lautzenheizer, "An iterative method for inverse scattering problems based on an exact gradient search", *Radio Science*, vol. 29, no. 4, pp. 1119–1130, July–August 1994.

[29] S. Caorsi, G. L. Gragnani, S. Medicina, M. Pastorino, and A. Pinto, "A Gibbs random fields-based active electromagnetic method for noninvasive diagnostics in biomedical applications", *Radio Science*, vol. 30, no. 1, pp. 291–301, January–February 1995.

[30] P. M. van den Berg and R. E. Kleinman, "A total variation enhanced modified gradient algorithm for profile reconstruction", *Inverse Problems*, vol. 11, pp. L5–L10, 1995.

[31] H. Carfantan, *Approche bayésienne pour un problème inverse non linéaire en imagerie à ondes diffractées*, PhD thesis, Université de Paris-Sud, Orsay, December 1996.

[32] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific, Belmont, Massachussetts, 1995.

[33] L. Garnero, A. Franchois, J.-P. Hugonin, C. Pichot, and N. Joachimowicz, "Microwave imaging – complex permittivity reconstruction by simulated annealing", *IEEE Transactions on Microwave Theory and Technology*, vol. 39, no. 11, pp. 1801–1807, November 1991.

[34] J. J. Xia, M. Habashy, Tarek, and J. A. Kong, "Profile inversion in a cylindrically stratified lossy medium", *Radio Science*, vol. 29, no. 4, pp. 1131–1141, July–August 1994.

[35] P. Lobel, R. Kleinman, C. Pichot, L. Blanc-Féraud, and M. Barlaud, "Conjugate gradient method for solving inverse scattering with experimental data", *IEEE Antennas and Propagation Magazine*, vol. 38, no. 3, pp. 48–51, June 1996.

[36] H. Carfantan and A. Mohammad-Djafari, "A Bayesian approach for nonlinear inverse scattering tomographic imaging", in *Proceedings of IEEE ICASSP*, Detroit, U.S.A., May 1995, vol. IV, pp. 2311–2314.

[37] H. Carfantan, A. Mohammad-Djafari, and J. Idier, "A single site update algorithm for nonlinear diffraction tomography", in *Proceedings of IEEE ICASSP*, Munich, Germany, 1997, pp. 2837–2840.

[38] M. Nikolova and A. Mohammad-Djafari, "Eddy current tomography using a binary Markov model", *Signal Processing*, vol. 49, pp. 119–132, 1996.